|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  2nd Meeting: Shanghai, CN, 13–19 Oct. 2012 | Document: JCT3V-B0092 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Removing temporal motion vector prediction for depth map coding** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Chih-Ming Fu, Yi-Wen Chen, Jian-Liang Lin, Yu-Wen Huang, and Shawmin Lei  No. 1, Dusing Rd. 1, Hsinchu Science Park, Hsinchu, Taiwan 30078 | Tel: Email: | Shawmin Lei +886-3-5670766 ext. 25555  {chihming.fu, yiwen.chen, jl.lin, yuwen.huang,shawmin.lei}@mediatek.com |
| *Source:* | MediaTek Inc. | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

In the HEVC-based 3D video coding, HTM-4.0.1, the depth coding requires motion information of temporal collocated pictures to derive temporal motion vector predictor (TMVP). In this proposal, TMVP and the related syntax are removed for depth coding to reduce encoder/decoder complexity and related motion information buffer. Experimental results showed that the average BD-rate change is 0.0% compared with the HTM-4.0.1. By using the proposed method, the encoder/decoder complexity is reduced and the related motion information buffer (about 1.4 mega bits for a HD sequence) for depth map can also be removed while keeping the same coding performance.

# Introduction

In the HEVC-based 3D video coding, HTM-4.0.1 [1], to support temporal motion vector predictor (TMVP) for depth coding, an extra buffer is required to store the motion information of all reference depth pictures. For each depth picture, the temporal motion information of a 64x64 block includes motion vector, reference picture index, and inter/intra mode information. For high definition video, the motion information buffer will become significant. We list the requirement of related information in Table 1.

**Table 1. The requirement of related information for temporal motion information**

|  |  |
| --- | --- |
| MV information | Number of bits |
| Motion vector | X: 16 bits, Y:16 bits |
| Reference index | 5×2 bits |
| Inter/Intra mode | 1 bit |

We can calculate the buffer requirement of MV buffer under common test conditions [2]. Total number of MV bits for depth map coding: 1920 × 1080 (maximum image size) / 4 / 4 / 16 (basic unit size of MV information) × 43 (number of bits for each MV unit) × 4 (number of reference frame) = 1.4 Mega bits. The buffer size required for TMVP is quite large for on chip memory. If we put this buffer to external memory, it may constraint the reconstruction speed. Moreover, deriving TMVP is more complex than any other spatial motion vector especially the temporal motion vector requiring scaling operation.

# Proposed Method and Experimental Results

We propose to disable TMVP for depth coding. Specifically, if sps\_temporal\_mvp\_enable\_flag is enabled and current coding picture is a depth map, all syntax related to TMVP (e.g. slice-level enable\_TMVP\_flag, **collocated\_from\_l0\_flag,** **collocated\_ref\_idx**) can be removed, so that the decoder can remove MV buffer in the worst-case. The proposed methods are integrated into HTM-4.0.1 [1] and all tests are conducted under common test conditions [2]. The experimental result is shown in Table 2. The results show that there is almost no change on coding performance with the proposed change.

**Table 2. The BD-rate result of removing TMVP in depth map coding**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Video 1 | Video 2 | Video only | Synthesized only | Coded & synthesized | Enc time | Dec time | Ren time |
| Balloons | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 96.9% | 99.2% | 101.3% |
| Kendo | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 99.2% | 100.7% | 100.3% |
| Newspapercc | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 99.9% | 99.5% | 99.9% |
| GhostTownFly | 0.0% | 0.0% | 0.0% | 0.1% | 0.1% | 98.4% | 99.7% | 101.2% |
| PoznanHall2 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 98.7% | 99.5% | 100.0% |
| PoznanStreet | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 98.2% | 99.4% | 101.0% |
| UndoDancer | 0.0% | 0.0% | 0.0% | 0.2% | 0.1% | 97.7% | 98.9% | 101.4% |
| 1024x768 | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 98.6% | 99.8% | 100.5% |
| 1920x1088 | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 98.3% | 99.4% | 100.9% |
| **average** | **0.0%** | **0.0%** | **0.0%** | **0.1%** | **0.0%** | **98.4%** | **99.6%** | **100.7%** |

# Conclusion

In this proposal, TMVP for depth coding is removed by removing the related syntax when coding a depth picture. Experimental results showed that the proposed method does not have any coding loss in terms of the average BD-rate for coded and synthesized views. By using the proposed method, the encoder/decoder complexity can be reduced and the related motion information buffer (about 1.4 mega bits) for depth coding can also be removed, which can reduce the worst-case of the decoder buffer requirement in hardware design.

# Patent rights declaration (s)

**MediaTek Inc. may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

# References

[1] HTM-4.0.1, <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-4.0.1/>

[2] Dmytro Rusanovskyy, Karsten Müller, Anthony Vetro, “Common Test Conditions of 3DV Core Experiments”, Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A1100, July 2012, Stockholm.